

#3

Attorney Docket: 3036/50059
PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: SIMON PAUL DAVIS ET AL.

Serial No.: 09/898,486 Group Art Unit:

Filed: JULY 5, 2001 Examiner:

Title: IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES

CLAIM FOR PRIORITY UNDER 35 U.S.C. §119

Commissioner for Patents
Washington, D.C. 20231

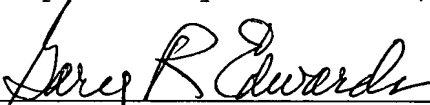
Sir:

The benefit of the filing date of prior foreign application numbers 0016476.4 and 0024737.9, filed in Great Britain on July 5, 2000 and October 10, 2000 respectively, is hereby requested and the right of priority under 35 U.S.C. §119 is hereby claimed.

In support of this claim, filed herewith is a certified copy of the original foreign application.

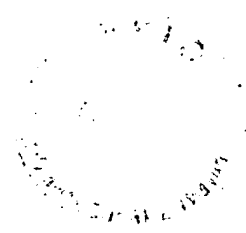
Respectfully submitted,

October 22, 2001



Gary R. Edwards
Registration No. 31,824

CROWELL & MORING, LLP
P.O. Box 14300
Washington, DC 20044-4300
Telephone No.: (202) 624-2500
Facsimile No.: (202) 628-8844



THIS PAGE BLANK (USPTO)



INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

**CERTIFIED COPY OF
PRIORITY DOCUMENT**



**CERTIFIED COPY OF
PRIORITY DOCUMENT**

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name in which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., p.l.c. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

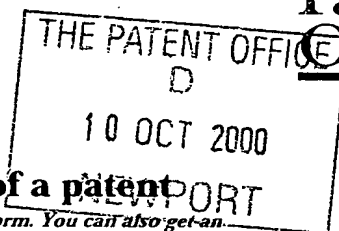
Signed

[Signature]

Dated 14 June 2001

**CERTIFIED COPY OF
PRIORITY DOCUMENT**

THIS PAGE BLANK (USPTO)



Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference		2000P04865/GB/R76/MM/rr	
2. Patent application number (The Patent Office will fill in this part)		0024737.9 10 OCT 2000	
3. Full name, address and postcode of the or of each applicant (underline all surnames)		ROKE MANOR RESEARCH LIMITED OLD SALISBURY LANE ROMSEY, HAMPSHIRE, SO51 0ZN UNITED KINGDOM	
Patents ADP number (if you know it)			
If the applicant is a corporate body, give the country/state of its incorporation		UNITED KINGDOM	
4. Title of the invention		IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES	
5. Name of your agent (if you have one)		DEREK ALLEN	
"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)		Siemens Shared Services Limited Intellectual Property Department Siemens House, Oldbury Bracknell, Berkshire RG12 8FZ United Kingdom	
Patents ADP number (if you know it)			
6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number	Country	Priority application number (if you know it)	Date of filing (day / month / year)
	GB	0016476.4	05/07/2000
7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application	Date of filing (day / month / year)	
8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:	YES		
a) any applicant named in part 3 is not an inventor, or			
b) there is an inventor who is not named as an applicant, or			
c) any named applicant is a corporate body.			
See note (d))			

Patents Form 1/77

- Enter the number of sheets for any of the following items you are filing with this form.
Do not count copies of the same document

Continuation sheets of this form	0
Description	11
Claim(s)	1
Abstract	1
Drawing(s)	3

238

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77) **3** /

Request for preliminary examination and search (Patents Form 9/77) **1** /

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

Margaret Mackett

Date 06.10.2000

Margaret Mackett

12. Name and daytime telephone number of person to contact in the United Kingdom
- | | |
|------------------|--------------|
| Margaret Mackett | 01344 396808 |
|------------------|--------------|

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES

The present invention relates to improvements in or relating to switching devices and is more particularly concerned with a technique for transmitting control information across a switching device.

Data is transferred over the Internet by means of a plurality of packet switches in accordance with a standard protocol known as Internet Protocol (IP). IP is a protocol based on the transfer of variable sized portions of data known as packets. All network traffic involves the transportation of packets of data. Packet switches are devices for accepting incoming packets; temporarily storing each packet; and then forwarding the packets to another part of the network. A packet switch receives packets of data on a plurality of input ports and transfers each packet to a specific one of a plurality of output ports. The packets of data can be of variable length or of fixed length. A packet switch may include a router, or routing device, or a circuit switch.

Traffic volume in the Internet is growing exponentially, almost doubling every 3 months, and the capacity of conventional IP routers is insufficient to meet this demand. There is thus an urgent requirement for products that can route IP traffic at extremely large aggregate bandwidths in the order of several terabits per second. Such routing devices are termed "terabit routers".

Terabit routers require a scalable high capacity communications path between the point at which packets arrive at the router (the "ingress") and the point at which the packets leave the router (the "egress").

The packets transferred in accordance with IP can (and do) vary in size. Within routers it has been found useful to pass data in fixed sized

units. In routers then data packets are partitioned into small fixed sized units, known as cells.

One suitable technique for implementing a scalable communications path is a backplane device, known as a cell based cross-bar. Data packets
5 are partitioned into cells by a plurality of ingress means for passage across the cross-bar.

The plurality of ingress means provide respective interfaces between incoming communications channels carrying incoming data and the cross-bar. Similarly a plurality of egress means provide respective interfaces
10 between the cross-bar and outgoing communications channels carrying outgoing data.

A general terabit router architecture bears some similarity to conventional router architecture. Packets of data arrive at input port(s) of ingress means and are routed as cells across the cross-bar to a
15 predetermined egress means which reassembles the packets and transmits them across its output port(s). Each ingress means maintains a separate packet queue for each egress means.

The ingress and egress means may be implemented as line interface cards (LICs). Since one of the line functions regularly undertaken by the
20 ingress and egress means is forwarding, LICs may also be known as 'forwarders'. Further functions include congestion control and maintenance of external interfaces, input ports and output ports.

In a conventional cell based cross-bar, each ingress means is connected to one or more of the egress means. However, each ingress
25 means is only capable of connecting to one egress means at any one time. Likewise, each egress means is only capable of connecting to one ingress means at a time.

All ingress means transmit in parallel and independently across the cross-bar. Furthermore, cell transmission is synchronised with a cell cycle, having a period of, for example, 108.8ns.

5 The ingress means simultaneously each transmit a new cell with each new cell cycle. The pattern of transmissions from the ingress means across the cross-bar to the egress means changes at the end of every cell cycle.

A cross-bar controller is provided for efficient allocation of the bandwidth across the cross-bar. It calculates the rates that each ingress means must transmit to each egress means. This is the same as the rate at
10 which data must be transmitted from each packet queue. The calculation makes use of real-time information, including traffic measurements and indications from the ingress means. The indications from the ingress means include monitoring the current rates, queue lengths and buffer full flags. The details of the calculation are discussed more rigorously in the
15 copending British Patent Application Number 9907313.2 (docket number F21558/98P4863).

The cross-bar controller performs a further task; it serves to schedule the transfer of data efficiently across the cross-bar whilst maintaining the calculated rates. At the end of each cell cycle, the cross-bar controller
20 communicates with the ingress and egress means as follows. First, the cross-bar controller calculates and transmits to each ingress means the identity of the next packet queue from which to transmit. Secondly, the cross-bar controller calculates and transmits to each egress means the identity of the ingress from which it must receive.

25 The architecture described above gives rise to two requirements:-

(i) the need for a means for each ingress means to transmit traffic measurements and indications to the cross-bar controller; and

(ii) the need for a means for the cross-bar controller to send configuration information to each ingress and each egress means.

It is possible to provide dedicated communications paths to meet these requirements. However such a solution requires additional hardware,
5 which is expensive in terms of increased power consumption, installation and materials.

It is therefore an object of the invention to obviate or at least mitigate the aforementioned problems.

In accordance with the present invention, there is provided a
10 switching device for user data in the form of cells, the switching device comprising:-

a backplane;

a plurality of ingress means connected to an input side of the backplane;

15 a plurality of egress means connected to an output side of the backplane;

for each ingress means, associated slicing means for converting cells into slices for transfer across the backplane;

20 for each egress means, associated de-slicing means for reforming the slices into cells; and

backplane control means for controlling the backplane in accordance with control slices which are interspersed with slices carrying the user data.

Advantageously, the control slices are spaced in time.

Preferably, the control slices are located in predetermined timeslots.

25 The present invention has the advantage that it is faster and more efficient as the use of slices rather than the significantly larger cells allows relatively low data rates for control channels without the consequent

increase in latency of control traffic that use of cells would impose. Also, it removes the need for separate control hardware for the backplane.

In one embodiment of the present invention, there is provided a router device having a plurality of ingress line function means, a plurality of egress line function means, a backplane and a controller means, wherein the transmission of signals from the plurality of ingress line function means to the controller means and signals from the controller means to each of the ingress line function means and each of the egress line function means takes place across the backplane.

For a better understanding of the present invention, reference will now be made, by way of example only, to the accompanying drawings in which:-

Figure 1 illustrates a terabit router architecture;

Figure 2 illustrates a switching device in accordance with the present invention; and

Figure 3 illustrates the operation of the slices in accordance with the present invention.

Figure 1 illustrates a conventional terabit router architecture 100 in which packets arrive at ingress forwarders 102, 104, 106 via their input port(s) (not shown) and are routed across a cross-bar 110 to a correct egress forwarder 120 which transmits them across its output port(s) (not shown). Each ingress forwarder 102, 104, 106 maintains a separate packet queue for each egress forwarder 120.

Ingress forwarder 102 has three queues q_{11} , q_{12} , q_{13} of data packets ready for transfer to three separate egress forwarders (only egress forwarder 120 being shown). Data in q_{11} is destined for egress forwarder 120 via the cross-bar 110. Similarly, three queues q_{21} , q_{22} , q_{23} and q_{31} , q_{32} , q_{33} are formed respectively in each of the ingress forwarders 104, 106. Although

three queues are shown in each ingress forwarder 102, it will be appreciated that any number of queues can be present in each ingress forwarder 102, 104, 106.

Generally speaking, each queue may be defined such that j represents the ingress, k represents the egress, and q_{jk} represents the packet queue at the ingress j for the packets destined for egress k .

It will be appreciated that although only one egress forwarder 120 is shown in Figure 1, the number of egress forwarders will normally be the same as the number of ingress forwarders.

By way of explanation, a cell based cross-bar is characterised as follows:

- a) Each ingress line function may be connected to any egress line functions.
- b) Each ingress line function may only be connected to one egress line function at a time.
- c) Each egress line function may only be connected to one ingress line function at a time.
- d) All ingress line functions transmit in parallel across the cross-bar.
- e) Data is transmitted across the cross-bar in small fixed sized cells, for example, a cell size is typically 64 octets.
- f) Cell transmission is synchronised across all the ingress line functions. This means that for each cell cycle, each ingress line function starts transmitting the next cell at the same time.
- g) The cross-bar is reconfigured at the end of every cell cycle.

As shown in Figure 1, packets of data arriving at the ingress forwarders 102, 104, 106 via their input port(s) (not shown) and are routed across the cross-bar 120 to the correct egress forwarders 120 which

transmits them across its output port(s) (also not shown). Each ingress forwarder 102, 104, 106 maintains a separate packet queue for each egress forwarder 120, for example q_{11} , q_{12} , q_{13} , q_{21} , q_{22} , q_{23} , q_{31} , q_{32} , q_{33} .

A cell based cross-bar arrangement 200 in accordance with the present invention is shown in Figure 2. The arrangement 200 comprises a plurality of ingress forwarders 210 and a plurality of egress forwarders 220 connected to a cross-bar or backplane 230. Here, each ingress forwarder 212, 214, 216, 218 may be connected to one or more of the egress forwarders 222, 224, 226, 228. However, as mentioned above, each ingress forwarder 212, 214, 216, 218 may only be connected to one egress forwarder 222, 224, 226, 228 at a time and each egress forwarder 222, 224, 226, 228 may only be connected to one ingress forwarder at a time 212, 214, 216, 218.

The cross-bar arrangement 200 is controlled by a cross-bar controller 240 which is physically connected to the backplane 230 via connection 232. The cross-bar controller 240 is also logically connected to each ingress forwarder 212, 214, 216, 218 via logical links 242, 244 and to each egress forwarder 222, 224, 226, 228 via logical link 246. The cross-bar controller 240 co-ordinates the transmission and reception of cells via links 242, 244, 246.

The term 'logical link' means that there is no physical connection between the cross-bar controller 240 and the ingress and egress forwarders 212, 214, 216, 218, 222, 224, 226, 228, and all transfer of control information either from or to the cross-bar controller 240 is made via the backplane 230.

Each ingress forwarder 212, 214, 216, 218 communicates traffic measurements and notifications for the use of the cross-bar controller 240, via logical link 242. The cross-bar controller 240 communicates to each

ingress forwarder 212, 214, 216, 218 which cell it is to send next, via logical link 244. The cross-bar controller 240 also communicates to each egress forwarder 222, 224, 226, 228 information indicating from which ingress forwarder 212, 214, 216, 218 to receive data, via logical link 246.

- 5 The cross-bar controller 240 allocates connections between ingress forwarders 212, 214, 216, 218 and egress forwarders 222, 224, 226, 228 and informs the respective forwarders accordingly for each cell cycle in turn.

- 10 In accordance with the present invention, the backplane 230 is configured such that data is transmitted thereacross in slices. A slice is a fixed size portion of a cell – typically each cell is subdivided into eight slices.

- Each ingress forwarder 212, 214, 216, 218 includes slicing means 252, 254, 256, 258 for dividing cells into slices for transmission across the backplane 230. Each egress forwarder 222, 224, 226, 228 includes de-slicing means 262, 264, 266, 268 for receiving slices from the backplane 230 re-forming the original cells. The backplane 230 deals only with slices and not cells.

- 20 Cells are input to ingress forwarders 212, 214, 216, 218, the cells are sliced in the slicing means 252, 254, 256, 258 and transmitted across the backplane 230 to de-slicing means 262, 264, 266, 268 in the egress forwarders 222, 224, 226, 228 and the output from each egress forwarder 222, 224, 226, 228 is in the form of cells.

- 25 The ingress and egress forwarders 212, 214, 216, 218, 222, 224, 226, 228 are synchronised so that they each send or receive slices simultaneously. At each slice time, each ingress forwarder 212, 214, 216, 218 will transmit a slice which can be received by one or more egress forwarders 222, 224, 226, 228. Likewise, at each slice time, each egress

forwarder 222, 224, 226, 228 can receive a slice from one and only one ingress forwarder 212, 214, 216, 218. Each egress forwarder 222, 224, 226, 228 is responsible for selecting the correct slice.

5 As the backplane 230 only operates on slices, the cross-bar controller 240 includes a slicing means 270 for providing control information in the form of slices. In accordance with the present invention, the control information from the cross-bar controller 240 is interleaved with user data across the backplane 230.

10 User data is conveyed across the backplane 230 as cells consisting of some fixed integral number of slices. This is described in more detail with reference to Figure 3.

In Figure 3, slice timeslot patterns 302, 304, 306, 308 for each of the ingress forwarders 212, 214, 216, 218 of Figure 2 are shown. Each slice timeslot pattern 302, 304, 306, 308 is different and comprises a control
15 slice timeslot 312, 314, 316, 318 for carrying control information from the associated ingress forwarder 212, 214, 216, 218 to the cross-bar controller 240, a control slice timeslot 322, 324, 326, 328 for carrying control information from the cross-bar controller 240 to each ingress forwarder 212, 214, 216, 218, and a control timeslot 332, 334, 336, 338 for carrying
20 control information from the cross-bar controller 240 to the egress forwarders 222, 224, 226, 228. As shown, for each ingress forwarder 212, 214, 216, 218, the position of its control slice timeslots 312, 314, 316, 318, 322, 324, 326, 328, 332, 334, 336, 338 is different to each other ingress forwarder.

25 Data to be transferred across the backplane 230 in the form of slices which are fitted into slice timeslots around the control slice timeslots. For example, if ingress forwarder 212 has eight data slices to transmit, it will place the first slice in the first timeslot before control slice timeslot 312, six

slices in the next six timeslots following the control slice timeslot 312 and the last slice in the timeslot following the control slice timeslot 322.

Similarly, for ingress forwarder 214 having eight data slices to transmit, the first three slices will be placed in the three timeslots prior to the control
5 slice timeslot 314 and the remaining five timeslots will be in the five timeslots following the control slice timeslot 314, and so on.

If ingress forwarder 216 has fifteen slices to transmit, then the first five slices are placed in the first five timeslots, the next six slices are placed in the six timeslots following the control slice timeslot 316, the next two
10 slices are placed in the two timeslots following the control slice timeslot 326, and the remaining two slices are placed in the two timeslots following the control slice timeslot 336. Similarly, for ingress forwarder 218 having fifteen slices to transmit, the first seven slices will be placed in the first seven timeslots, the next six slices will be placed in the six timeslots
15 following the control slice timeslot 318, and the last two slices will be placed in the two timeslots between the control slice timeslots 328 and 338.

For transmission of control information from ingress forwarders 212, 214, 216, 218 to the cross-bar controller 240, each ingress forwarder 212, 214, 216, 218 is assigned a dedicated slice timeslot which it uses to send
20 information to the controller 240. The timeslots do not overlap. When the timeslot assigned to a given ingress forwarder 212, 214, 216, 218 is reached, that ingress forwarder transmits a slice of control information, interrupting its transmission of user data. The cross-bar controller 240 selects the ingress forwarder 212, 214, 216, 218 from which to receive
25 control information according to the current timeslot number.

When receiving user data from a given ingress forwarder 212, 214, 216, 218, an egress forwarder 222, 224, 226, 228 ignores information in a slice timeslot if that timeslot is assigned to the given ingress forwarder for

transmission of control information. The position of the control slice timeslot is determined by fixed global information, for example, the position of a forwarder in a physical rack of forwarders or LICs. This makes it simple for each forwarder to determine which slice timeslot is used by each forwarder for this purpose.

For transmission of control information from the cross-bar controller 240 to ingress and egress forwarders 212, 214, 216, 218, 222, 224, 226, 228, the same technique is used except that each forwarder is assigned a dedicated timeslot on which to receive.

Where the backplane 230 supports broadcast traffic, that is, the transmission of information to all ingress and/or egress forwarders simultaneously, this can be achieved by using a single control slice timeslot. All recipients would receive information using this timeslot. Such a control slice timeslot may be in addition to the control slice timeslots 312, 314, 316, 318, 322, 324, 326, 328, 332, 334, 336, 338, or it may replace one or more of such timeslots in accordance with a particular application.

It will be readily understood that although the preceding discussion has been in terms of optical terabit routers, the apparatus of the present invention are capable of implementation in a wide variety of routing devices, including switches and routers, and that these routing devices can be either purely electronic, part electronic/part optical or optical in nature.

CLAIMS:

1. A switching device for user data in the form of cells, the switching device comprising:-
 - a backplane;
 - a plurality of ingress means connected to an input side of the backplane;
 - a plurality of egress means connected to an output side of the backplane;
 - for each ingress means, associated slicing means for converting cells into slices for transfer across the backplane;
 - for each egress means, associated de-slicing means for reforming the slices into cells; and
 - backplane control means for controlling the backplane in accordance with control slices which are interspersed with the user data.
2. A device according to claim 1, wherein the control slices are spaced in time.
3. A device according to claim 1 or 2, wherein the control slices are located in predetermined timeslots.
4. A switching device substantially as hereinbefore described with reference to Figures 2 and 3 of the accompanying drawings.

ABSTRACT
IMPROVEMENTS IN OR RELATING TO ROUTING DEVICES

The invention relates to a router device (200) having a plurality of ingress line interface cards (LICs) (212, 214, 216, 218), a plurality of egress LICs (222, 224, 226, 228)), a backplane (230) and a controller (240). Transmission of signals from the ingress LICs (212, 214, 216, 218) to the controller (240) and likewise the transmission of signals from the controller (240) to each of the ingress LICs (212, 214, 216, 218) and each of the egress LICs (222, 224, 226, 228) takes place across the backplane (230). Each ingress LIC (212, 214, 216, 218) is provided with a dedicated timeslot in which it can send information to the controller (240) via connection (242). Information is sent in a slice within the dedicated timeslot and each egress LIC (222, 224, 226, 228) ignores data sent by a given ingress LIC within the timeslot assigned to said ingress LIC. A similar system is used for transmission of communications from the controller (240) to the LICs. It is thus possible to avoid provision of additional, dedicated communications paths between the LICs (ingress and egress) and the controller (240).

(Fig. 2)

THIS PAGE BLANK (USPTO)

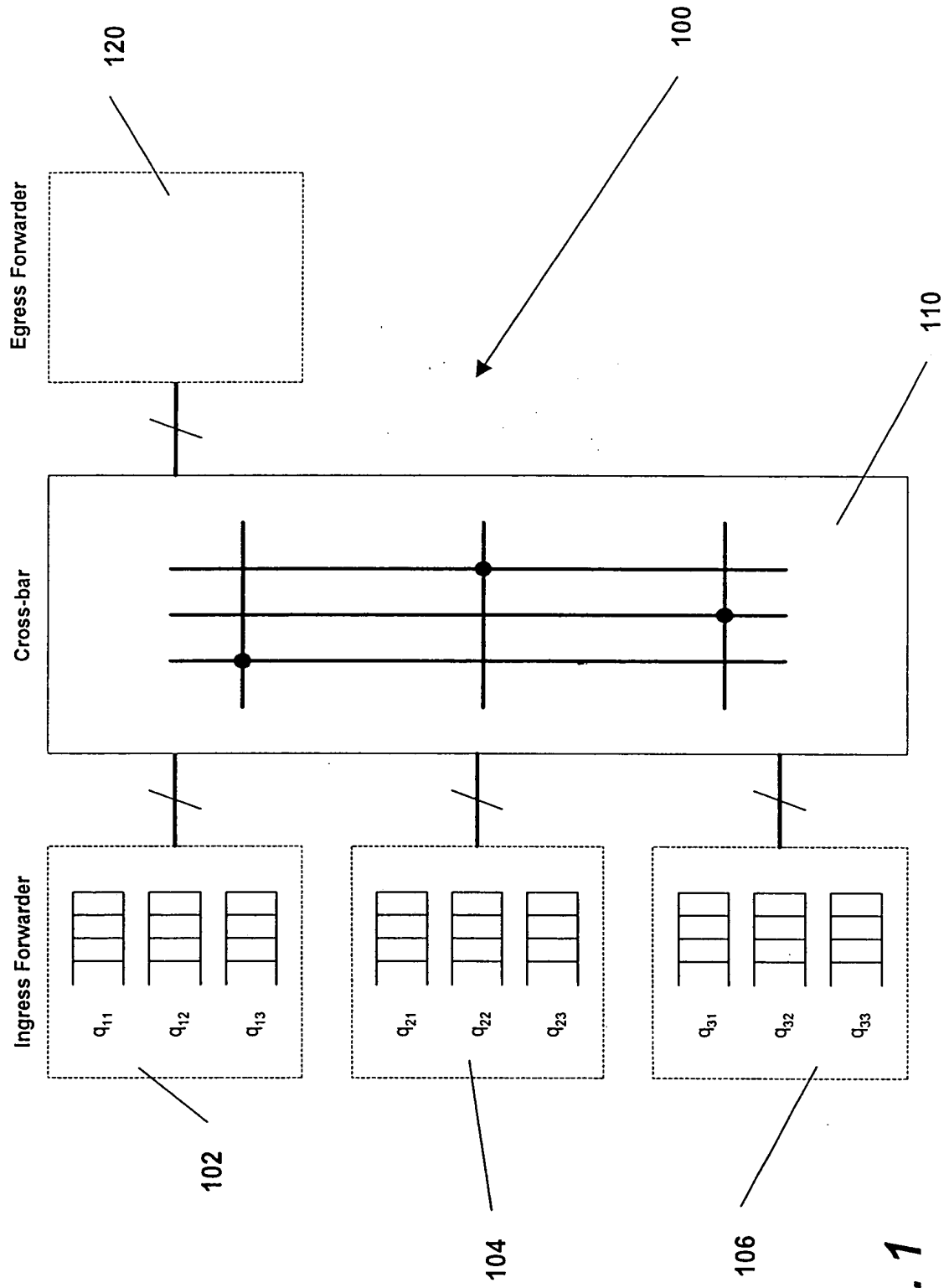


Fig. 1

THIS PAGE BLANK (USPTO)

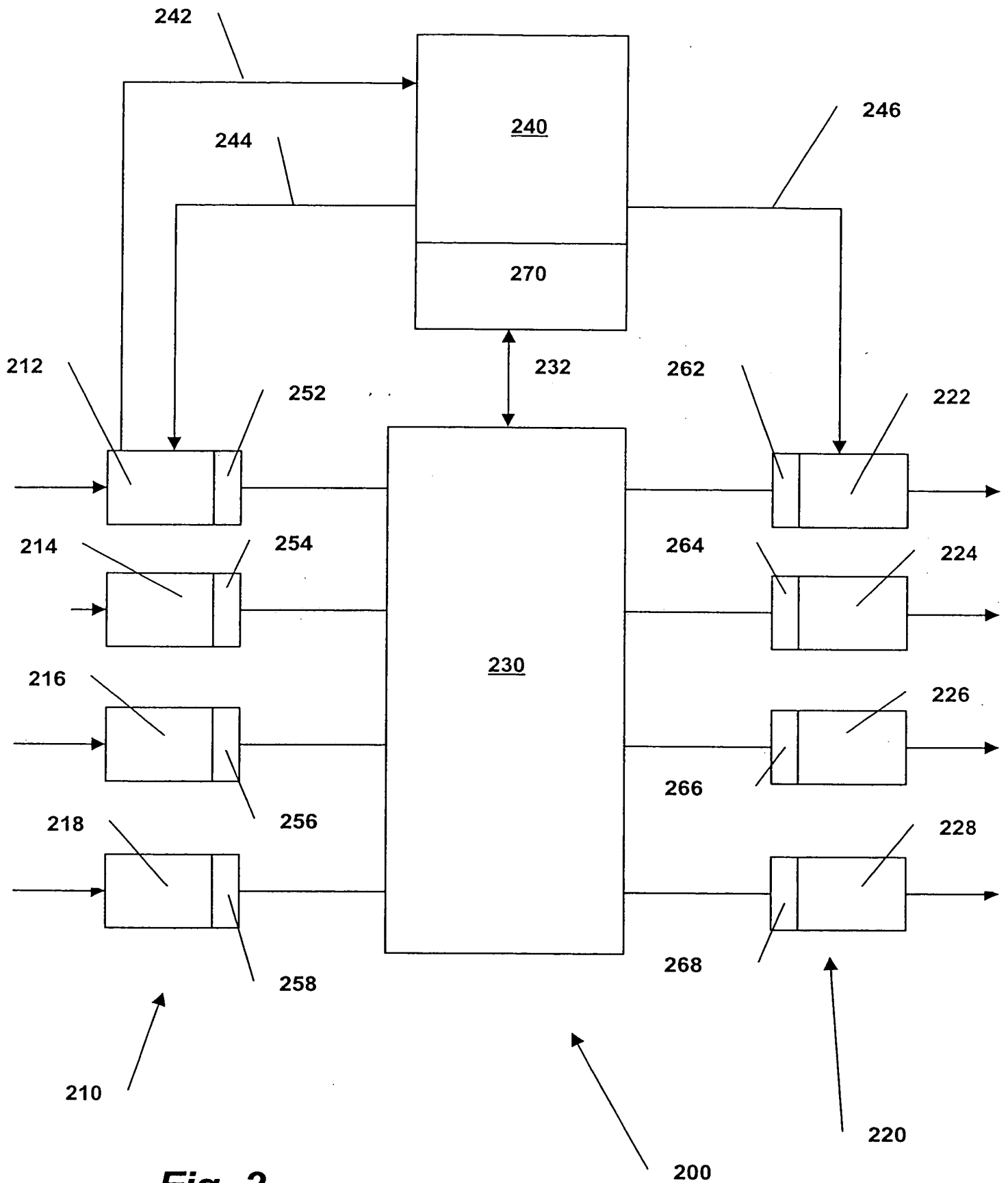


Fig. 2

THIS PAGE BLANK (USPTO)

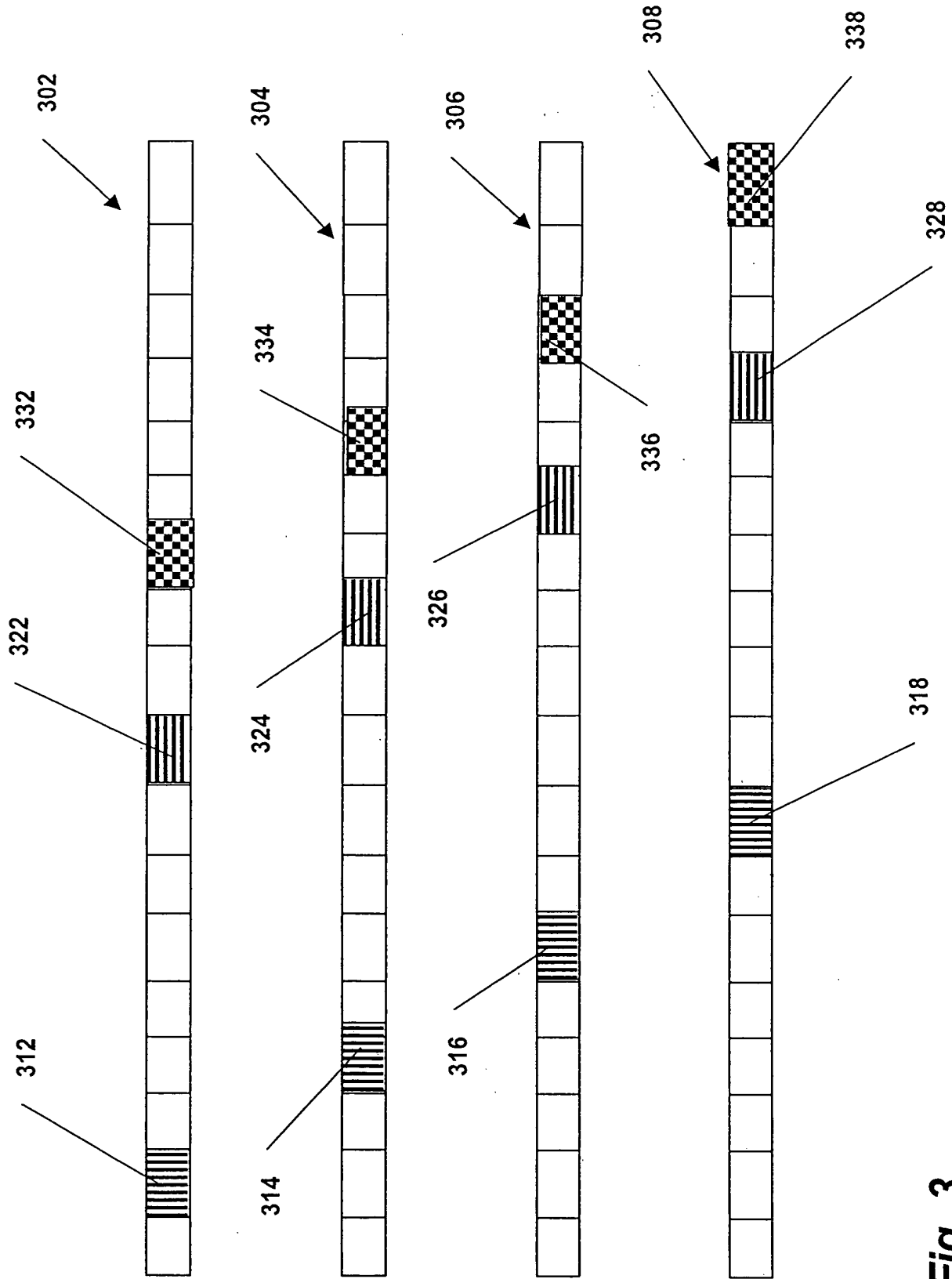


Fig. 3

THIS PAGE BLANK (USPTO)